

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-208945

(43)Date of publication of application : 26.07.2002

(51)Int.Cl.

H04L 12/56

(21)Application number : 2001-003000

(71)Applicant : FUJITSU LTD

(22)Date of filing : 10.01.2001

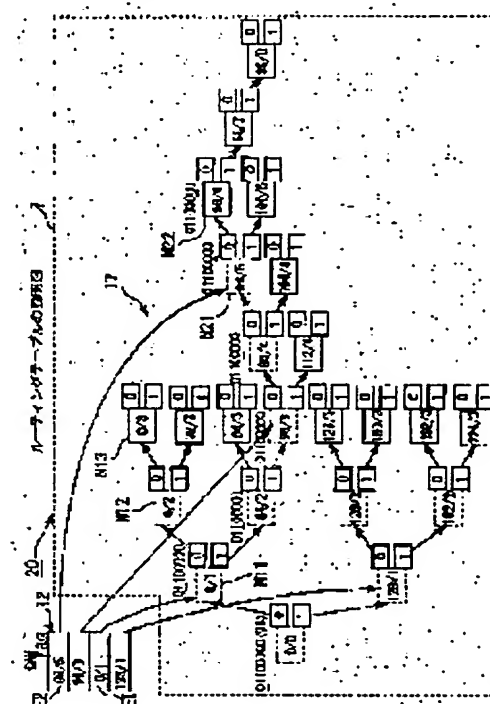
(72)Inventor : EZAKI YUTAKA
KAWASAKI TAKESHI

(54) DESTINATION INFORMATION MANAGING APPARATUS FOR PACKET

(57)Abstract:

PROBLEM TO BE SOLVED: To provide a destination information managing apparatus for packet capable of keeping more prefixes than those of a CAM, and retrieving prefixes more rapidly than a RAM.

SOLUTION: The apparatus for managing the destination information of the packets for selecting the paths of the packets is provided with a first memory, a second memory having a retrieval speed lower than the first memory and the degree of integration higher than the first memory, and a compiling section for compiling a path table storing the destination information so that the information is over the first and second memories.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号
特開2002-208945
(P 2 0 0 2 - 2 0 8 9 4 5 A)
(43) 公開日 平成14年7月26日 (2002.7.26)

(51) Int. Cl. ⁷	識別記号	F I	テーマコード (参考)
H04L 12/56	100	H04L 12/56	100 Z 5K030

審査請求 未請求 請求項の数5 O L (全18頁)

(21) 出願番号 特願2001-3000 (P 2001-3000)

(22) 出願日 平成13年1月10日 (2001.1.10)

(71) 出願人 000005223

富士通株式会社

神奈川県川崎市中原区上小田中4丁目1番1号

(72) 発明者 江▲崎▼ 裕

神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

(72) 発明者 川崎 健

神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

(74) 代理人 100089244

弁理士 遠山 勉 (外1名)

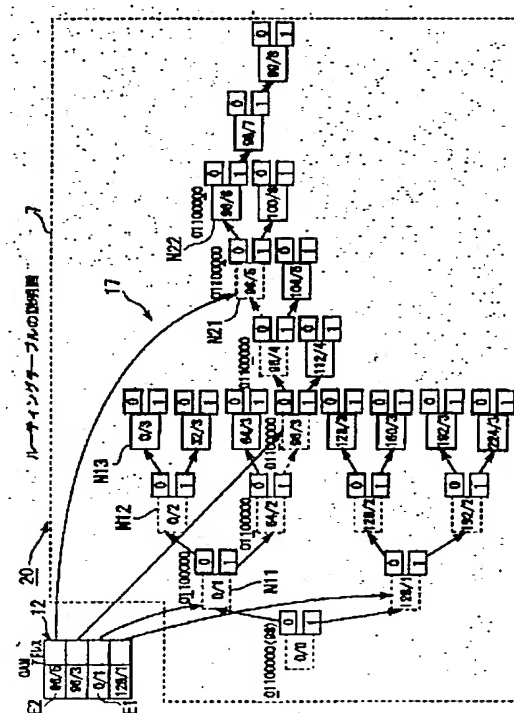
最終頁に続く

(54) 【発明の名称】 パケットの宛先情報管理装置

(57) 【要約】

【課題】 CAMよりも多くのプレフィックスを保持することができ、且つRAMよりも高速にプレフィックスを検索することが可能なパケットの宛先情報管理装置を提供する。

【解決手段】 パケットの経路を選択するためのパケットの宛先情報を管理する装置であって、第1メモリと、この第1メモリよりも情報の検索速度が遅く且つ集積度が高い第2メモリと、前記宛先情報が前記第1メモリと前記第2メモリとに跨る状態で格納された経路表を作成する作成部とを備える。



【特許請求の範囲】

【請求項 1】パケットの経路を選択するためのパケットの宛先情報を管理する装置であって、

第 1 メモリと、

前記第 1 のメモリより情報の検索速度が遅く且つ集積度が高い第 2 メモリと、

前記宛先情報が前記第 1 メモリと前記第 2 メモリとに跨がる状態で格納された経路表を作成する作成部と、を備えたパケットの宛先情報管理装置。

【請求項 2】前記第 1 メモリの使用率の閾値を記憶した 10 閾値記憶部をさらに備え、

前記作成部は、前記第 1 メモリの使用率を取得し、取得した使用率が前記閾値記憶部に記憶された閾値未満である場合には、前記経路表に格納すべき宛先情報を前記第 1 メモリのみに格納する請求項 1 記載のパケットの宛先情報管理装置。

【請求項 3】前記経路表に格納された宛先情報に対応する出力ルートの情報を得る場合の最大許容時間を記憶した許容時間記憶部をさらに備え、

前記作成部は、前記最大許容時間内で前記経路表から前記パケットの宛先情報が検索され、検索された宛先情報 20 に対応する出力ルートの情報が取得されるように、前記宛先情報を前記第 1 メモリと前記第 2 メモリとに跨がる状態で格納する請求項 1 又は 2 記載のパケットの宛先情報管理装置。

【請求項 4】前記最大許容時間内で可能な前記第 2 メモリへの最大アクセス回数を記憶したアクセス回数記憶部をさらに備え、

前記作成部は、前記経路表に格納された宛先情報の検索に必要な前記第 2 メモリへのアクセス回数が前記最大アクセス回数以上となる場合には、前記アクセス回数が前記最大アクセス回数未満となるように、前記第 2 メモリに格納された、又は格納される予定の宛先情報を前記第 1 メモリに格納する請求項 3 記載のパケットの宛先情報管理装置。

【請求項 5】パケットの宛先情報が第 1 メモリとこの第 1 のメモリよりも情報の検索速度が遅く且つ集積度が高い第 2 メモリとに跨る状態で格納された経路表と、入力されたパケットに対応する宛先情報を前記経路表から検索し、検索した宛先情報に対応する出力ルートの情報 40 を取得する検索部と、を備えたパケットの経路選択装置。

【発明の詳細な説明】

【0 0 0 1】

【発明の属する技術分野】本発明はパケットの宛先情報管理装置に関し、例えば、インターネットやイントラネット 50 で利用されるインターネットプロトコル(I P)アドレスを管理する装置に関する。

【0 0 0 2】

【従来の技術】インターネットやイントラネットは、T 50

C P / I P を利用した通信方式であり、コンピュータ間の簡易な通信方式としてローカル・エリア・ネットワーク(L A N)からワイド・エリア・ネットワーク(W A N)まで広く利用されている。I P では、あらゆる情報がパケットに区切られて転送される。パケットは、数十バイト〜数十キロバイト程度の可変長のサイズを持つ。パケットには、送信先アドレス(ディスティネーションアドレス)、送信元アドレス(ソースアドレス)等を含むヘッダが添付される。

【0 0 0 3】インターネットは、一般に、ルータと呼ばれる通信装置を相互接続することによって構築される。ルータは、パケットの宛先情報に基づいて、パケットの経路を選択する(ルーティング: フォワーディングともいう)。即ち、ルータは、パケットを受信すると、受信したパケットのヘッダに格納された宛先 I P アドレスに対応する次ホップアドレス(パケットを次に受信すべきルータの I P アドレス)をルーティングテーブル(「フォワーディングテーブル」ともいう)から取得する。ルータは、次ホップアドレスを取得すると、当該次ホップアドレスを持つルータへパケットを転送する。ルーティングテーブルは、パケットの宛先のネットワーク(「ドメイン」と呼ばれる)の I P アドレス(「プレフィクス」と呼ばれる)と、プレフィクスに対応する次ホップアドレスとを記憶している。

【0 0 0 4】ダイナミックルーティングが採用される場合、ルータは、ルーティングプロトコルに従って、ルーティング情報(プレフィクス等)を他のルータと相互に交換し、ルーティングテーブルを作成する。ルーティングテーブルは、上記したルーティング情報の交換が定期的、周期的に実行されることによって、数分〜数時間毎に更新される。ルーティングプロトコルには、例えば、O S P F (Open Shortest Path First)、R I P (Routing Information Protocol)、B G P (Border Gateway Protocol)等がある。

【0 0 0 5】ルーティングは、ルータに到着したパケット毎に実行される。このため、ルーティング性能は、ルータの性能を決定する大きな要因となっている。インターネットサービスプロバイダ(I S P)によって使用されるルータでは、約 2 5 万の I P アドレスに対応する次ホップアドレスを検索可能なルーティングテーブルが要求される。このような大容量の情報を高速に検索可能であり、且つ効率的に保存する技術が従来から検討されてきた。

【0 0 0 6】ルーティングテーブルの構成方法には、大略して、C A M (Contents Accessable Memory: 「連想メモリ」と呼ばれる)などの専用のメモリを利用する方法(「C A M 方式」という)と、R A M (Random Access Memory)を用いてツリーを構成する方法(「R A M ツリー方式」という)とがある。

【0 0 0 7】I P アドレスは、ネットワークアドレス

(プレフィクス)とホストアドレスとで構成される。有効なプレフィクス長(有効長)は、ネットワーク構成やアドレス情報の集約(アグリゲーション)によって変化する。具体的に説明すると、一般に、ネットワークは、コア(中継)、エッジ等からなる階層で構成されており、ルータが保持すべき機能やルーティング情報の細かさ(アドレスのアグリゲーションの細かさ)は、ルータ毎に異なる。例えば、ネットワークが、図12に示すように構成されている場合、LANでは、ルーティングの際に、IPアドレスの全アドレス長(32ビット)がプレフィクス(／32)として使用される。これに対し、アクセスネットワークでは、ルーティングの際に、IPアドレスの上位24ビット(／24)がプレフィクスとして使用される。さらに、コアネットワークでは、ルーティングの際に、IPアドレスの上位16ビット(／16)がプレフィクスとして使用される。

【0008】このように、ネットワークの中心部に進むに従ってルーティングに使用されるプレフィクスが短くなる。これに従い、LANとアクセスネットワークとを結ぶルータ(エッジルータ)は、IPアドレスの全バイト又は上位24バイトを利用してルーティングを行う。また、アクセスネットワークとコアネットワークとを結ぶルータ(コアルータ)は、IPアドレスの上位24バイト又は上位16バイトを利用してルーティングを行う。

【0009】但し、実際には、CIDR(Classless Inter-Domain Routing)によって、個々のルータが受け取ったルーティング情報に基づく経路集約の計算が行われる。このため、プレフィクス長は固定ではなく、ドメインのIPアドレス毎に異なるプレフィクス長が用いられる。なお、CIDRのネットワークアドレスは、IPアドレス(プレフィクスに相当)とマスク値とで表現される。マスクは、ネットワークアドレスの有効ビットであり、アドレスの最左端(最上位ビット)からの可変の連続ビット列で表現される。

【0010】ところで、ルータに入力されるバケットには、有効長が明示されていない。このため、ルータは、バケットの宛先のIPアドレスに対応するプレフィクスを検索する場合には、ロングストマッチ(longest match)検索と呼ばれる特別な検索手順を実施する。ロングストマッチ検索は、宛先のIPアドレスのうち、プレフィクスを構成するビット列について、最左端(最上位)からのビット列が最長に一致するプレフィクスのエントリをルーティングテーブルから検索する方法である。

【0011】CAMは、エントリとして記憶されている内容をキーとして、キーに対応するエントリのアドレスを出力する特殊なメモリである。CAM方式を用いれば、宛先のIPアドレスに対応するプレフィクスを、RAMツリー方式に比べて高速に検索することができる。

【0012】RAMツリー方式では、RAM上にプレフィクスのエントリがツリー状に並べられ、宛先IPアド

レスの各ビットが最上位ビットから最下位ビットへ向かってビット毎に比較されることによって、対応するプレフィクスのエントリが検索される。図13は、RAMツリー方式による検索方法の説明図である。図13には、ツリー構成方法の1つとしてのパトリシア(PATRICIA)ツリーの論理構成が示されている。

【0013】パトリシアツリーでは、エントリデータとして、プレフィクス値、マスク値(有効長)、0ポインタ(ゼロポインタ)、1ポインタ(ワンポインタ)、及びルーティング情報(次ホップアドレス)が、RAM上に設定される(図14参照)。パトリシアツリーは、エントリデータに応じて設けられた複数のノードN及びリンクLからなる。各ノードNは、上記したプレフィクス値と有効長とを有し、さらに、ノードNでの処理対象となるビット(0又は1)に対応するポインタP1及びポインタP2を有している。ポインタP1又はポインタP2のジャンプ先に相当するノードNがある場合には、ノードN同士はリンクLで接続される。

【0014】図13には、“000xxxxx(0/3)”, “001xxxx(32/3)”, “010xxxxx(64/3)”, “011000xx(96/6)”, “01100011(99/8)”, “011001xx(100/6)”, “01101xxx(104/5)”, “0111xxxx(112/4)”, “100xxxxx(128/3)”, “101xxxxx(160/3)”, “110xxxxx(192/3)”, 及び “111xxxxx(224/3)” の各エントリデータに応じたパトリシアツリーが示されている。ここでは、例として、プレフィクスの最大有効長が、8ビットで設定されている。

【0015】エントリデータ中の“x”は、プレフィクス中の無効ビットを示す。括弧中の左側の数字は、プレフィクス値を示し、エントリデータ中の“x”に“0”を代入した場合における2進数の値を10進数に換算した値である。また、右側の数字は、有効長を示す。例えば、エントリデータ“001xxxxx”は、プレフィクス“32”を示し、有効長は3ビットである。

【0016】バケットの宛先IPアドレスのプレフィクスが、図13に示すように“01100010(98)”である場合には、最初のノードN(図13の最も左側のノードN)において、最左端のビット(最上位ビット)が0か1かが判定される。このとき、ビットは“0”であるので、ポインタP1に従った次のノード(プレフィクス値が“0”で有効長が1のノードN((0/1)のノードN))へ進み、次のビットが0か1かが判定される。このような処理がポインタP1又はP2で指定されるリンク先が無くなるまで繰り返行われる(図13における太線の矢印参照)。その後、リンク先が或るノードNで無くなった時に、宛先のIPアドレスに対応するプレフィクス(宛先のネットワークアドレス)が特定される。そして、特定されたプレフィクスに対応する次ホップアドレスが出力ルートの情報としてルーティングテーブルから取り出される。

【0017】実際のRAM上では、パトリシアツリー

10

20

30

40

50

は、図14に示すようなツリーテーブルに格納される。例えば、図14に示すツリーテーブルを用い、“010”というプレフィクスを検索する場合には、以下の処理が行われる。即ち、最初のエントリでは、最上位ビット“0”に基づいて0ポインタが調べられ、0ポインタで示されたエントリにジャンプする(S101)。次のエントリでは、次のビット“1”に基づいて1ポインタが調べられ、1ポインタで示された次のエントリにジャンプする(S102)。そして、次のエントリでは、最下位ビット“0”に基づいて0ポインタが調べられる。このとき、0ポインタには、ジャンプ先が無いことを示すエンドフラグがジャンプ先のアドレスの代わりに格納されている。このため、当該エントリに格納されている出力ルートの情報(次ホップアドレス)がツリーテーブルから読み出される(S103)。

【0018】

【発明が解決しようとする課題】RAMツリー方式では、プロセッサ(CPU、MPU等)が、ソフトウェアの実行によって、プレフィクスを構成するビット毎にRAM上のツリーテーブルを次々と読み込んでいく。このため、プロセッサの読み込み回数(メモリアクセスの回数)が有効長に比例する。IPアドレス(最長32ビット)の場合、有効長の平均値は20ビット程度である。このため、20回程度のメモリアクセスが必要になる。このことにより、例えば、1回のRAMアクセスを30ナノ秒[ns]程度で実行可能なRAMを使用しても、検索が終了するまでに平均600[ns]、遅い場合には1マイクロ秒[μs]程度かかる。従って、RAMツリー方式では、パケットの転送性能を1Mpps[mega packet per second]程度までしか上げることができなかった。

【0019】RAMは集積度が高く、バトリシアツリーは、比較的効率的にプレフィクスのエントリを格納することができる。このため、RAMツリー方式がISPのルータに採用された場合でも、ルーティングテーブルには必要なエントリを格納することができる。半導体技術の発達により、プロセッサがツリーテーブルからルーティング情報を検索する手順をハードウェア化した装置(“レイヤ3スイッチ”と呼ばれる)もある。

【0020】しかしながら、ツリー方式は、CAM方式に比べて検索速度が遅いという問題があった。具体的には、CAM方式が、数十メガppsでパケットを転送可能であるのに対し、RAMツリー方式では約1メガppsである。但し、CAMの集積度は、RAMに比べて低く、現状では、数千エントリ程度しか実現されていない。

【0021】上述したように、ルーティングテーブルに格納すべきエントリ数は、インターネットの拡大に伴って増加している。また、パケットの転送速度の向上も望まれている。このため、数百キロエントリという大容量で且つ数十メガppsという高速に検索可能なルーティ

ングテーブルを実現することが望まれている。

【0022】本発明の目的は、CAM方式で作成されたルーティングテーブルよりも多くのプレフィクスを保持することができ、且つRAMツリー方式で作成されたルーティングテーブルよりも高速にプレフィクスを検索することが可能なパケットの宛先情報管理装置を提供することである。

【0023】

【課題を解決するための手段】本発明は、上述した目的を達成するために以下の構成を採用する。

【0024】即ち、本発明は、パケットの経路を選択するためのパケットの宛先情報を管理する装置であって、第1メモリと、前記第1メモリよりも情報の検索速度が遅く且つ集積度が高い第2メモリと、前記宛先情報が前記第1メモリと前記第2メモリとに跨がる状態で格納された経路表を作成する作成部と、を備える。

【0025】本発明によると、第2メモリよりも検索速度が速い第1メモリに宛先情報の一部が格納されるので、第2メモリのみを用いて格納された宛先情報を検索する場合よりも速く宛先情報を検索することができる。また、第2メモリは、第1メモリよりも集積度が高いので、第1メモリのみを用いる場合よりも多くの宛先情報を格納することができる。

【0026】第1メモリは、例えばCAM(連想メモリ)であり、第2メモリは、例えばRAMである。宛先情報は、例えばパケットの宛先のIPネットワークアドレス(プレフィクス)や、VPN識別子とIPネットワークアドレスとの組み合わせである。

【0027】本発明は、前記第1メモリの使用率の閾値を記憶した閾値記憶部にさらに備え、前記作成部は、前記第1メモリの使用率を取得し、取得した使用率が前記閾値記憶部に記憶された閾値未満である場合には、前記経路表に格納すべき宛先情報を前記第1メモリのみに格納する、構成とするのが好ましい。

【0028】また、本発明は、パケットの宛先情報が第1メモリとこの第1のメモリよりも情報の検索速度が遅く且つ集積度が高い第2メモリとに跨る状態で格納された経路表と、入力されたパケットに対応する宛先情報を前記経路表から検索し、検索した宛先情報に対応する出力ルートの情報を取得する検索部と、を備えた経路選択装置である。

【0029】

【発明の実施の形態】以下、図面を参照して本発明の実施形態を説明する。なお、本発明は、実施形態の構成に限定されるものではない。

【0030】〔第1実施形態〕

〔ルータの構成〕図1は、本発明による宛先情報管理装置及び経路選択装置が適用された通信装置としてのルータ1の構成例を示す図である。ルータ1は、大きく分けて、物理層(レイヤ1)及びデータリンク層(レイヤ2)に

関する処理を実行する複数のインターフェイスカード 2 と、パケットのルーティング(ネットワーク層:レイヤ 3)に関する処理を行うルーティング機構 3 とを備えている。インターフェイスカード 2 は、パケットの出力ルート(入力ルート)毎に設けられている。

【0031】ルーティング機構 3 は、ルーティングプロトコル処理部 4 と、ルート計算・経路表作成部 5 (以下、「経路表作成部 5」と表記)と、ルーティングテーブル(経路表) 20 と、ヘッダ抽出部 8 と、アドレス検索部 9 と、転送処理部 10 とを備えている。なお、ルーティング機構 3 が、本発明の管理装置及び経路選択装置に相当し、経路表作成部 5 が本発明の作成部に相当し、アドレス検索部 9 が本発明の検索部に相当する。

【0032】ルーティングプロトコル処理部 4 は、RIP、OSPF、BGP 等の所定のルーティングプロトコルに従ってルーティング情報の交換処理を実行し、他のルータからルーティング情報(プレフィクス等)を取得し、取得したプレフィクスを経路表作成部 5 に与える。

【0033】経路表作成部 5 は、ルーティングプロトコル処理部 4 から受け取ったプレフィクスに対応する最適な出力ルートを計算し、プレフィクスと算出された出力ルートの情報(次ホップアドレス)とが格納されたルーティングテーブル 20 を作成する(本発明の閾値記憶部、許容時間記憶部、アクセス回数記憶部に相当)。

【0034】ルーティングテーブル 20 は、第 1 メモリとしての CAM 6 と、第 2 メモリとしての RAM 7 とを含んでいる。CAM 6 及び RAM 7 は、経路表作成部 5 によって格納されたプレフィクス及び次ホップアドレスに関するエントリを保持する。なお、本実施形態では、ダイナミックルーティングによってプレフィクスが取得される例について述べているが、本発明は、スタティックルーティングによって作成されたルーティングテーブルを持つルータにも適用可能である。

【0035】ヘッダ抽出部 8 は、各インターフェイスカード 2 からルーティング機構 3 に入力されたパケットのヘッダを抽出し、抽出したヘッダから宛先 IP アドレス(ディスティネーション IP アドレス)を取り出し、アドレス検索部 9 に与える。

【0036】アドレス検索部 9 は、ヘッダ抽出部 8 から受け取った宛先 IP アドレスのプレフィクスに対応する次ホップアドレスをルーティングテーブル 20 から読み出し、転送処理部 10 に与える。

【0037】転送処理部 10 は、ルーティング機構 3 に入力されたパケットをアドレス検索部 9 から受け取った次ホップアドレスに対応するインターフェイスカード 2 に入力する。

【0038】インターフェイスカード 2 は、ルーティング機構 3 から入力されたパケットに対し、レイヤ 1 及びレイヤ 2 に関する処理(例えば、宛先 MAC アドレスの書換)を行う。その後、インターフェイスカード 2 は、

パケットを次ホップアドレスに相当する他のルータへ転送する。

【0039】なお、ルーティングプロトコル処理部 4、経路表作成部 5、ヘッダ抽出部 8、アドレス検索部 9 は、CPU 等のプロセッサ(図示せず)がプログラムを実行することによって実現される機能である。転送処理部 9 は、例えば、ソフトウェア又はハードウェアで実現されるスイッチを用いて構成される。

【0040】〈CAM の構成〉次に、CAM 6 について説明する。CAM は、与えられたパターンにマッチするエントリのアドレスを出力する特殊なメモリである。比較パターンとして“0”又は“1”しか設定できない CAM は、“binary CAM”と呼ばれ、比較パターンとして“0”、“1”及び“d.c. (マスク)”の 3 パターンを設定可能な CAM は、“ternary CAM”と呼ばれる。

【0041】図 2 は、CAM 6 の構成例を示す図である。CAM 6 は、ビット毎に比較器を有する特殊な記憶素子を複数アレイ上に並べて構成されたデータアレイ及びマスクアレイ 12 と、ロングストマッチ検索やアドレスの優先制御を実行するプライオリティ・エンコーダ(priority encoder) 13 と、書込データ等を出入力するための入出力制御装置(I/O: 図示せず)とを備えており、CAM 6 の出力信号は、補助 RAM 14 に入力される。

【0042】ロングストマッチ検索の前提として、書込データとしてのプレフィクスとマスク情報との組み合わせが、I/O を通じてデータアレイ/マスクアレイ 12 にエントリとして書き込まれる。比較対象(出力ルートの検索対象)としての宛先 IP アドレスが I/O を通じて CAM 6 に入力されると、当該宛先 IP アドレスとデータアレイ/マスクアレイ 12 の各エントリとが比較され、比較対象の宛先 IP アドレスに一致するプレフィクスを保持したエントリのアドレス(CAM アドレス)が検出される。

【0043】プライオリティ・エンコーダ 13 は、複数の CAM アドレスが検出された場合に、宛先 IP アドレスに最長に一致する(有効長が最も長い) CAM アドレス、或いは、優先すべき CAM アドレスを特定し、特定された CAM アドレスが出力される。

【0044】ternary CAM としての CAM 6 によるプレフィクス検索の例を図 2 を用いて説明する。今、CAM 6 のデータアレイ/マスクアレイ 12 には、図 2 に示すように、複数のプレフィクスとマスク情報との組み合わせがエントリとして書き込まれている。この例では、エントリに保持されたプレフィクスは、有効長が長い順でソートされている。

【0045】ここで、例えば、或るパケットの宛先 IP アドレス“192.168.0.177”が CAM 11 に入力されたとする。すると、宛先 IP アドレス“192.168.0.177”と、データアレイ/マスクア

レイ 12 に保持されたエントリ (プレフィクスとマスク情報との組み合わせ) の夫々とが比較され、両者が合致する CAM アドレスを示す信号が、プライオリティ・エンコーダ 13 に入力される。図 2 に示す例では、CAM アドレス “1”, “1003”, “1007” 及び “65535” が、プライオリティ・エンコーダ 13 に入力される。

【0046】プライオリティ・エンコーダ 13 は、入力された信号のうち、アドレス値が最も若い CAM アドレス “1” を示す信号を選択し、CAM アドレス “1” を補助 RAM 14 に入力する。これによって、CAM アドレス “1” に保持された “192. 168. 0. 177 / 32” が、CAM 6 に入力された宛先 IP アドレスに対応するプレフィクスとして特定される。

【0047】補助 RAM 14 は、CAM アドレスに対応する次ホップアドレスを保持しており、CAM アドレス “1” が入力されると、入力された CAM アドレスに対応する次ホップアドレスを出力する。出力された次ホップアドレスは、アドレス検索部 9 に与えられる。その後、次ホップアドレスは、転送処理部 10 に与えられ、転送処理部 10 は、パケットを次ホップアドレスに対応するインターフェイスカード 2 に与える。インターフェイスカード 2 は、受け取ったパケットを次ホップへ転送する。

【0048】CAM 6 は、宛先 IP アドレスが CAM 11 に入力されてから CAM アドレスを示す信号が補助 RAM 14 に格納されるまでの動作 (宛先 IP アドレスに対するロングストマッチ検索) を一度に行うことができる。現状では、CAM は、5 千万 ~ 1 億サーチ / 秒 (= 10 ~ 20 ns / サーチ) の速度で高速な検索を行うことができる。但し、CAM 6 の集積度は、前述した比較機能を有するために、数 Mbit であり、RAM に比べて数分の 1 に落ちる。

【0049】〈RAM の構成〉RAM 7 には、図 13 に示したようなパトリシアツリー (ツリーテーブル) が作成される。即ち、RAM 7 は、プレフィクス、マスク長、0 ポインタ、1 ポインタ、及び次ホップアドレスをエントリとして格納したツリーテーブルを格納している (図 14 参照)。RAM ツリー及びツリーテーブルの構造自体は、図 13 及び図 14 に示したものと同一である。経路表作成部 5 は、新規のプレフィクスと、このプレフィクスに対応する次ホップアドレスを含むエントリをパトリシアツリー (ツリーテーブル) に追加する。この場合には、経路表作成部 5 は、図 3 に示す 4 つのパターンの何れかによってエントリを追加する。4 つのパターンは、以下の通りである。

(A) 第 1 パターン: 既存ノードに一致する場合

(B) 第 2 パターン: 既存ノードに子ノードとして追加する場合

(C) 第 3 パターン: 既存ノードの中間に追加する場合

(D) 第 4 パターン: 既存ノードから分木ノードとともに追加する場合

第 1 パターンでは、エントリは、ツリー上に既に存在するノード (図 3 (A) のノード N1) に追加される。第 2 パターンでは、エントリは、ツリー上に既に存在するノード (図 3 (B) のノード N2) の子ノード (図 3 (B) のノード N3) として追加される。第 3 パターンでは、エントリは、或る親ノード (図 3 (C) のノード N4) と子ノード (図 3 (C) のノード N5) との間の中間ノード (図 3 (C) のノード N6) として追加される。

【0050】第 4 パターンでは、エントリは、或る親ノード (図 3 (D) のノード N7) と子ノード (図 3 (D) のノード N8) との間の分木ノード (図 3 (D) のノード N9) とともに、分木ノードの子ノード (図 3 (D) のノード N10) として追加される。エントリが追加された場合、経路表作成部 5 は、追加されたエントリ (ノード) に関連するノードのポインタ (0 ポインタ及び / 又は 1 ポインタ) を正しく付け替える。これによって、RAM ツリーが修正される。

【0051】〈経路表を用いたアドレス検索〉次に、アドレス検索部 9 がルーティングテーブル 20 を用いて次ホップアドレスを検索する処理について説明する。図 4 は、実施形態におけるルーティングテーブル 20 の構成図である。図 4 において、CAM 6 は、ternary CAM を用いて構成されている。RAM 7 は、パトリシアツリー (RAM ツリー) 17 (のツリーテーブル) を格納している。

【0052】本実施形態では、宛先 IP アドレスに対応するプレフィクスは、以下の 3 つのパターンの何れかによってルーティングテーブル 20 に格納されている。

(A) CAM 6 のエントリとして格納

(B) RAM ツリー 17 のエントリとして格納

(C) CAM 6 及び RAM ツリー 17 に亘って (跨って) 格納

上記したパターン (C) では、プレフィクスが所定数の有効ビットからなることに鑑み、プレフィクスのうち、最左端のビット (最上位ビット) から所定ビットまで (第 1 の部分) を検索するためのエントリが CAM 16 にセットされ、上記所定ビットから最後のビット (最下位ビット) まで (第 2 の部分) を検索する為のエントリが、RAM ツリー 17 にセットされる。

【0053】例えば、プレフィクス “0000xxxx” (プレフィクス値 0 / 有効長 4) を、パターン (C) でルーティングテーブル 20 にセットする場合には、CAM 6 及び RAM ツリー 17 は、以下の構成を持つ。即ち、CAM 6 のデータアレイ / マスクアレイ 12 には、“0 (0xxxxxx) / 1” のエントリ E1 が設けられる。一方、RAM ツリー 17 には、“0 / 1” のエントリを保持したノード N11 が設けられる。

【0054】エントリ E1 のアドレス (CAM アドレス)

は、ノードN11のRAMアドレスにリンクされている。これによって、アドレス検索部9(図1参照)は、CAM16から出力されるエントリE1のアドレスを用いてノードN11にアクセスすることができる。

【0055】即ち、パターン(C)でプレフィクスがセットされる場合には、CAMアドレスに対応するRAMアドレス(RAMツリーのエントリ(ノード)のアドレス)が、CAMアドレスに対応する次ホップアドレスの代わりに、補助RAM14(図2参照)に格納される。ノードN11は、“0/2”のエントリを保持したノードN12に0ポインタを介してリンクされており、ノードN12は、“0/3”のエントリを保持したノードN13に0ポインタを介してリンクされている。

【0056】また、例えば、プレフィクス“0110000x”(プレフィクス値96/有効長7)について、CAM6及びRAMツリー17は以下の構成を持つ。即ち、CAM6のデータアレイ/マスクアレイ12には、“96(01100xx)/5”のエントリE2が設けられている。一方、RAMツリー17には、“96/5”のエントリを保持したノードN21が設けられている。

【0057】エントリE2のCAMアドレスは、ノードN21のRAMアドレスと補助RAM14を介してリンクされており、アドレス検索部9は、CAM16から出力されるエントリE2のCAMアドレスを用いてノードN21にアクセスすることができる。ノードN21は、“96/6”のエントリを保持したノードN22に0ポインタを介してリンクされている。

【0058】アドレス検索部9は、ヘッダ抽出部8から宛先IPアドレスを受け取ると、以下のルーティング処理を実行する。即ち、アドレス検索部9は、宛先IPアドレスをCAM6に入力する。ここで、入力された宛先IPアドレスが、例えば“01100000”である場合には、CAM6は、ロングストマッチ検索によって、プレフィクスの最上位ビットからの各ビット値が最長に一致する“96/5”のエントリE2のCAMアドレスを出力する。出力されたCAMアドレスは、補助RAM14に入力される。すると、エントリE2のCAMアドレスに対応するRAMアドレス(ノードN21のアドレス)が補助RAM14からアドレス検索部9に与えられる。

【0059】アドレス検索部9は、補助RAM14からRAMアドレスを受け取ると、受け取ったRAMアドレスを用いてRAMツリー17のノードN21にアクセスする。次に、アドレス検索部9は、ノードN21の0ポインタによって指定されたノードN22にアクセスする。

【0060】アドレス検索部9は、ノードN22にて0ポインタを参照する。0ポインタは、指定されるリンク先の代わりに、リンク先がないことを示すエンドフラグを格納している。そこで、アドレス検索部9は、宛先IPアドレスに対応するプレフィクスが“011000xx(96

/6)”であるものとして、プレフィクス“96/6”に対応する次ホップアドレスをRAM7から読み出す。このようにして、宛先IPアドレスに対応するプレフィクスが特定され、出力ルート(次ホップ)が決定される。

【0061】上述したルーティング処理によると、宛先IPアドレスに対応するプレフィクスのうち、最上位ビットから5ビット目までが、CAM16によって一度に検索される。その後、プレフィクスの6ビット目と7ビット目についての処理が、RAMツリー17を用いて行われ、宛先IPアドレスのプレフィクスが特定される。

【0062】上記ルーティング処理において、最上位ビットから5ビット目までに対する処理は、RAMツリー17のみを用いて行くと、5回のRAMアクセスを要する。これに対し、上記ルーティング処理は、CAM16を用いて一度に行うため、検索時間が短縮される。

【0063】また、7ビットの有効長を持つプレフィクスがRAMツリー17のみを用いて検索される場合には、7回のRAMアクセスが必要である。これに対し、上記ルーティング処理は、RAMアクセスを2回に抑えることができる。従って、検索に必要な時間が、RAMツリー17のみを用いてプレフィクスを検索する場合に比べて短縮される。一方、プレフィクスのエントリが、上記したパターン(B)によってRAMツリー17に設定される。このため、CAM6に設定されるエントリ数を減らすことができる。

【0064】従って、上述したルーティングテーブル20によれば、ルーティングテーブルがCAMのみを用いて構成される場合よりも多くのエントリを格納することができる。一方、ルーティングテーブル20によれば、1つのパケットについてルーティングに要する時間を、ルーティングテーブルがRAMツリーのみを用いて構成される場合よりも短縮することができ、ルーティング性能の向上を図ることができる。

【0065】〈エントリの追加〉ネットワークのトポロジは、ネットワークの再構築、ネットワークの保守・点検、リンクやノードの障害等によって変化する。トポロジが変化すると、ルータ1から或るネットワークへ至る最適なルートが変化する。ルータ1は、トポロジの変化に対応すべく、定期的又は周期的にルーティングテーブル20を更新する。即ち、ルータ1のルーティングプロトコル処理部4は、定期的又は周期的にルーティングプロトコルに従ってルーティング情報の交換を行い、他のルータから取得したルーティング情報からプレフィクスを抽出し、経路表作成部5に与える。経路表作成部5は、新たなプレフィクスをルーティングプロトコル処理部4から受け取った場合には、受け取ったプレフィクスへパケットを転送するための最適な出力ルートを計算し、計算結果に基づいて次ホップを決定する。

【0066】そして、経路表作成部5は、受け取ったプレフィクスをルーティングテーブル20にエントリとし

て格納し、格納された新たなエン트리と決定された次ホップアドレスとを関連づける。これによって、アドレス検索部 9 が、新たなエントリに対応する宛先 IP アドレスを持つパケットについて、対応する次ホップアドレスを得ることができる。

【0067】図 5 及び図 6 は、経路表作成部 5 によるエントリの追加処理を示すフローチャートである。図 5 において、ルータ 1 の管理者は、CAM 6 の利用率の閾値 $\eta_c(\max)$ と、RAM 7 の利用率の閾値 $\eta_r(\max)$ とを、経路表作成部 5 の図示せぬメモリに格納する(ステップ S 1)。

【0068】CAM 6 の利用率 η_c は、CAM 6 (のデータアレイ/マスクアレイ 12) に格納されたエントリ数を CAM 6 に格納可能な全エントリ数で除した値(CAM 6 のエントリ利用率)である。一方、RAM 7 の利用率 η_r は、RAM 7 に RAM ツリー 17 を構成するノードとして格納されたエントリ数を RAM 7 に格納可能な全エントリ数で除した値(RAM 7 のエントリ利用率)である。

【0069】次に、ルータ 1 の管理者は、1 回のルーティングの許容時間 T_{\max} と、RAM 7 へのアクセス時間 τ_r と、1 回のルーティングにおける RAM 7 への最大アクセス回数 $\delta_{\max} = T_{\max} / \tau_r$ とを、経路表作成部 5 の図示せぬメモリに格納する(ステップ S 2)。なお、ステップ S 1 及びステップ S 2 の処理は、ルータ 1 で自動的に行われるようにしても良い。

【0070】その後、ステップ S 3 以降の処理は、経路表作成部 5 が、ルーティングプロトコル処理部 4 からプレフィクスを受け取った場合に開始される。経路表作成部 5 は、ルーティングプロトコル処理部 4 から 1 以上のプレフィクスを受け取ると(ステップ S 3 ; Yes)、ルーティングテーブル 20 に追加すべき新たなプレフィクスが受け取ったプレフィクスに含まれているか否かを判定し(ステップ S 4)、含まれていない場合には処理をステップ S 3 に戻し、含まれている場合には、処理をステップ S 5 に進める。

【0071】ステップ S 5 では、経路表作成部 5 は、追加対象の新規のプレフィクス(新たに得られたプレフィクス)を“N”に設定し、プレフィクス“N”の有効長を“n”に設定する。続いて、経路表作成部 5 は、プレフィクス“N”を用いてルーティングテーブル 20 を検索し、ルーティングテーブル 20 にエントリとして格納されているプレフィクスのうち、プレフィクス“N”と最長に一致するプレフィクス“M”と、プレフィクス“M”の有効長“m”を求める(ステップ S 6)。次に、有効長“n”から有効長“m”を減算し、有効長の差分 δ を求める(ステップ S 7)。

【0072】次に、経路表作成部 5 は、図 6 に示すように、CAM 6 の利用率 η_c を計測するとともに、RAM 7 の利用率 η_r を計測する(ステップ S 8)。次に、経路

表作成部 5 は、CAM 6 に十分な空き領域があるか否かを判定する(ステップ S 9)。即ち、経路表作成部 5 は、CAM 6 の利用率 η_c が条件“ $\eta_c < \eta_c(\max)$ ”を満たすか否かを判定する。

【0073】このとき、条件が満たされる場合(ステップ S 9 ; Yes)には、経路表作成部 5 は、プレフィクス“N”のエントリを CAM 6 に追加(格納)する(ステップ S 10)。

【0074】このように、利用率 η_c の値が閾値 $\eta_c(\max)$ 未満である場合には、CAM 6 の利用率 η_c が低く、CAM 6 に十分な空き領域があるものとして、プレフィクス“N”のエントリが CAM 6 のみに格納される。これによって、ルーティング時間を短縮することができる。これに対し、条件が満たされない場合には、処理がステップ S 11 に進む。

【0075】ステップ S 11 では、経路表作成部 5 は、RAM 7 に十分な空き領域があるか否かを判定する。即ち、経路表作成部 5 は、RAM 7 の利用率 η_r が条件“ $\eta_r < \eta_r(\max)$ ”を満たすか否かを判定する。

【0076】このとき、条件が満たされない場合(ステップ S 11 ; No)には、ルータ 1 は、プレフィクス“N”のエントリをルーティングテーブル 20 に登録できないものとしてエラーを出力する(ステップ S 12)。これに対し、条件が満たされる場合(ステップ S 11 ; Yes)には、処理がステップ S 13 に進む。

【0077】ステップ S 13 では、経路表作成部 5 は、差分 δ に対応する RAM ツリーが作成された場合に、作成された RAM ツリーにアクセスする回数(追加ツリーの段数)が RAM 7 の最大アクセス回数 δ_{\max} 未満か否か(許容範囲内か否か)を判定する。

【0078】即ち、経路表作成部 5 は、差分 δ が条件“ $\delta_{\max} > \delta$ ”を満たすか否かを判定する。このとき、経路表作成部 5 は、差分 δ が条件を満たす場合(ステップ S 13 ; Yes)には、差分 δ に対応する各ビットのノード(エントリ)を RAM 7 に登録する(ステップ S 15)。

【0079】ここで、ルーティングテーブル 20 には、既に、プレフィクス“M”が格納されている。このとき、プレフィクス“M”が CAM 6 のみに格納されている場合には、差分 δ に対応する各ビットのノードのうち、最もルート(root)側(上流側)のノードと CAM 6 に格納されたプレフィクス“M”のエントリとが、上記手法によって、リンクで接続される。

【0080】これに対し、プレフィクス“M”の有効ビットの最下位ビットが RAM 7 にノードとして格納されている場合には、差分 δ に対応する各ビットのノードは、当該ノードの後段にリンク付けされて配置される。

【0081】これらによって、プレフィクス“N”の有効ビットのうち、最上位ビットから所定ビット(第 1 の部分)までが CAM 6 に格納され、所定ビットから最下

位ビットまで(第2の部分)がRAMツリー17に格納される。即ち、プレフィクス“N”が、CAM6をルート(root)とするツリーに格納される。これによって、CAM6のオーバーフローが防止される。

【0082】また、プレフィクス“M”に相当するプレフィクスがルーティングテーブル20に格納されていない場合には、プレフィクス“N”は、ステップS15において、RAM7のみに格納される。なお、CAM6が“0/0”のエントリを有していれば、ルーティングテーブル20に格納される全てのプレフィクスは、CAM6とRAMツリー17とに跨って格納される。

【0083】経路表作成部5は、差分 δ が条件を満たさない場合(ステップS13; No)には、CAM6のエントリの書換処理を行う(ステップS14)。即ち、経路表作成部5は、CAM6に格納されているCAMエントリの何れかを削除する。続いて、経路表作成部5は、プレフィクス“N”をCAM6にエントリとして格納する。或いは、経路表作成部5は、プレフィクス“N”のうち、最上位ビットから所定ビットまでをCAM6に格納し、且つ所定ビットから最下位ビットまでに対応するノード(エントリ)をRAM7に格納(追加)する。

【0084】上記処理によると、プレフィクス“N”に関するRAMツリー17の段数が減少する。従って、アドレス検索部9によるRAM7のアクセス回数が減少する。このため、アドレス検索部9がプレフィクス“N”に対応する次ホップアドレスを宛先IPアドレスに対応する次ホップアドレスとして取得するために必要な時間が許容時間 T_{max} を超えてしまうことが防止される。

【0085】経路表作成部5は、CAM6からCAMエントリを削除する場合には、少なくとも以下の方法の何れかをを用いて削除すべきCAMエントリを選択する。

- (1)ランダムでCAMエントリを選択
- (2)有効長が短い順でCAMエントリを選択
- (3)有効長が長い順でCAMエントリを選択
- (4)リンクされたRAMツリーが短い順でCAMエントリを選択

上記(3)の方法を使用する場合には、経路表作成部5は、ホストアドレスのエントリである(/32)のCAMエントリを選択することができる。上記(4)の方法を使用する場合には、CAMエントリとリンクで接続されているRAMエントリからのRAMツリー長を比較し、RAMツリー長が短い方のCAMエントリを削除する。

【0086】例えば、図8に示す例では、経路表作成部5は、プレフィクス“96/3”のRAMエントリのノードを起点とするRAMツリーと、プレフィクス“96/5”のRAMエントリのノードを起点とするRAMツリーとを比較し、“96/5”のRAMエントリのノードの方がRAMツリーが短いため、当該RAMエントリに対応する“96/5”のCAMエントリを削除する。

【0087】経路表作成部5は、上記(4)の方法を用い

た選択処理を実施するため、RAMエントリ上、又は指示せぬ作業メモリの対応する場所に、CAMエントリとリンクされたRAMツリーの段数を書き込む。そして、経路表作成部5は、上記(4)の方法でCAMエントリを削除する場合に、予め書き込んだ段数を参照し、最も短いRAMツリーとリンクで接続されたCAMエントリを選択して削除する。

【0088】図7は、CAM6にエントリを追加することでRAMツリー17の段数を減らす処理例を示す図である。図7では、CAM6AのエントリA'(符号25)と、ノード26(エントリA)、ノード27(エントリB)、ノード28(エントリM)、及びノード29(エントリX)とからなるRAMツリー17Aとによってプレフィクス“M”が登録されている。このようなルーティングテーブル20Aに対し、新規のプレフィクス“N”を登録する場合を考える。

【0089】この場合において、例えば、以下のように仮定する。即ち、ノード29(エントリX)の後段にノード30(エントリY)を追加すれば、プレフィクス“N”をルーティングテーブル20Aに登録することができる。しかし、ノード30が追加されると、RAMツリー17Aの段数(RAMへのアクセス回数)が最大アクセス回数 δ_{max} を超えてしまう。

【0090】この場合、経路表作成部5は、プレフィクス“N”についてのRAMツリー17Aのルート(root)であるノード26(エントリA)と、追加すべきリーフ(leaf)のノード30(エントリY)との間に位置するノード32(エントリM)に対応するエントリM'(符号31)をCAM6Aに追加する(ステップS14)。その後、ノード30(エントリY)を追加する(ステップS15)。

【0091】このようにすれば、ノード30が追加されても、プレフィクス“N”に係るRAMツリー17Aの段数は、ノード30の追加前のほぼ半分になる。これによって、RAMへのアクセス回数が最大アクセス回数 δ_{max} 未満に抑えられる。また、RAMへのアクセス回数が減少するので、プレフィクス“N”の検索時間が短縮される。

【0092】なお、図7を用いてRAMツリーの段数をほぼ半分にする例について説明したが、当該例のようにルートのノード26からリーフのノード30までに亘って分木がない場合には、ノード29のエントリXやノード30のエントリYに対応するエントリをCAM6Aに格納すれば、RAMへのアクセス回数をさらに減らすことができる。

【0093】このように、CAMへのエントリの追加によってRAMツリーの段数(アクセス回数)を減らす場合には、他のプレフィクスの検索に影響を与えない範囲で、可能な限りアクセス回数が減るようなエントリを選択するのが好ましい。

【0094】図6に戻って、ステップS15の後、経路

成部5は、CAMエントリ“96/3”をルート(root)とするRAMツリーの全てが最大アークセア回数 δ_{max} を超えないことをチェックする。この場合、アライクヌ“96/3”のRAMエントリのノードからアライクヌ“99/8”のRAMエントリのノードに到るRAMツリーの段数を数え、段数が最大アークセア回数 δ_{max} を超えないことを確認する。

【0101】(RAMエントリの削除) 次に、経路表作成部5によるRAMエントリの削除処理の例を示す図である。図9は、RAMエントリの削除処理の例を示す図である。図9において、経路表作成部5は、アライクヌ“99/8”のRAMエントリ(ノード32)を、削除しようとしている。 “99/8”のRAMエントリのCA M6上の起点は、アライクヌ“96/4”のエントリ(符号38)である。また、図9中の各ノード33、34、35のRAMエントリは実体のない作業用のエントリであり、各ノード32、36、37のRAMエントリは、実体を有するRAMエントリである。

【0102】最初に、経路表作成部5は、ノード32に子ノードがないことを確認する。次に、経路表作成部5は、ノード32の親ノードに相当する“96/6”のノード33を参照する。ノード33の子ノードはノード32のみであるので、経路表作成部5は、ノード33(“96/6”のRAMエントリ)を削除可能と判断する。

【0103】次に、経路表作成部5は、ノード33の親ノードに相当するノード34を参照する。ノード34の子ノードはノード33のみであるので、経路表作成部5は、ノード34(“96/5”のエントリ)も削除可能と判断する。

【0104】次に、経路表作成部5は、ノード34の親ノードに相当するノード35を参照する。ノード35の子ノードはノード34のみであるので、経路表作成部5は、ノード35(“96/4”のエントリ)を削除可能と判断する。また、ノード35は、CAMエントリ38とリンクしているので、CAMエントリ38も削除可能と判断する。

【0105】次に、経路表作成部5は、ノード35の親ノードに相当するノード36を参照する。ノード36は、ノード34の他にノード37(“112/4”のエントリ)を子ノードとするので、経路表作成部5は、ノード36(“96/3”のエントリ)を削除できないと判断する。

【0106】すると、経路表作成部5は、削除可能と判断したノード32～35及びCAMエントリ38を削除する。このようにして、不要なRAMエントリ及びCAMエントリが削除される。

【0107】上記した削除処理を経路表作成部5が実行するためには、或るRAMエントリ(ノード)がCAMエントリとリンクされているか否かの情報が予め保持され

表作成部5は、ルーチングテーブル20に格納されたアライクヌ“N”を、対応する次ホップアライクヌと関連づける。

【0095】即ち、対応する次ホップアライクヌが既にRAM7に格納されている場合には、両者は、この次ホップアライクヌがアライクヌ“N”に対応する次ホップアライクヌとしてアライクヌ検索部9に与えられるように関連づけられる。このとき、対応する次ホップアライクヌがこの次ホップアライクヌと当該アライクヌ“N”とが関連づけられた状態でルーチングテーブル20に格納される。

【0096】これによって、アライクヌ“N”を宛先とするパケットがルーチング機構3に入力された場合には、アライクヌ検索部9がアライクヌ“N”に対応する次ホップアライクヌを取得し、パケットが次ホップアライクヌへ転送される。上記処理が終了すると、処理がステップS4へ戻り、経路表作成部5は、ルーチングテーブル20から受け取った他のアライクヌについて上記したステップS5～ステップS15の処理を行う。

【0097】(CAMエントリの削除) 経路表作成部5は、ルーチングテーブル20の更新処理において、CAM6(又はCAM6A)のエントリが不足した場合に、強制的に何れかのCAMエントリを削除する。以下、CAMエントリの削除処理について説明する。

【0098】CAM6のエントリが強制的に削除される場合には、CAMエントリが削除された後に実施されるアライクヌの検索に支障がないことが確認されなければならない。このため、経路表作成部5は、或るCAMエントリを強制的に削除する場合には、削除しようとするCAMエントリのCAMアライクヌで指し示されるRAMアライクヌを起点とするRAMツリーが他のCAMエントリ(例えば、削除対象のCAMエントリ)に保持されたアライクヌよりも短いアライクヌを保持したCAMエントリ)によって検索することができ、且つ当該他のCAMエントリとリンクされているノードから数えた上記RAMツリーの段数が最大アークセア回数 δ_{max} を超えないことを確認する。

【0099】図8は、強制的なCAMエントリの削除処理の例を示す図である。図8に示すように、CAM6Bは、“96/3”及び“96/5”のCAMエントリを記憶しており、経路表作成部5が、例えば“96/5”のCAMエントリを削除すると仮定する。この場合、経路表作成部5は、以下の手順により削除可能なことを確認する。

【0100】即ち、経路表作成部5は、削除しようとするCAMエントリ“96/5”より短いアライクヌのCAMエントリをCAM6Bから探す。この例では、CAMエントリ“96/3”が見つかる。次に、経路表作

ていることを要する。このため、各ノードのRAMエントリには、CAMエントリとのリンクの有無を示すビットが用意され、CAMエントリとのリンクがある場合には、“0”と“1”との一方が設定され、リンクがない場合には、“0”と“1”との他方が設定される。

【0108】経路表作成部5は、上記削除処理において、削除か否かの判断の対象となっているRAMエントリに設定されたビットを参照し、リンクが設定されていることを示すビットを有するRAMエントリを削除する場合には、当該RAMエントリにリンクされているCAMエントリも削除する。

【0109】第1実施形態によると、ルーティングテーブル20がCAM6とRAM7とを有し、プレフィクスがCAM6とRAM7とに跨って格納される。このため、RAMツリーのみを用いたルーティングに比べてルーティングに要する時間を短縮することができる。また、CAM6とRAM7とを併用することで、CAM6飲みを用いる場合に比べて多くのプレフィクス(のエントリ)をルーティングテーブル20に格納することができる。

【0110】さらに、経路表作成部5が、CAM6の利用率 η_r 、RAMの利用率 η_c 、差分 δ 、最大アクセス回数 δ_{\max} 、RAMツリー1回のアクセス時間 τ_r 、パケットのルーティングの最大許容時間 T_{\max} を用いて新規のプレフィクスの登録処理(ルーティングテーブル20の更新処理)を行う。これによって、プレフィクスを最適な状態でルーティングテーブル20に格納することができる。このため、CAM6のオーバーフローを防止しつつ、RAMへのアクセス回数(RAMツリー17の段数)を可能な限り減らすことでルーティングの高速化を図ることができる。

【0111】このため、高速且つ大容量のルーティングテーブルを構成することができる。

【0112】〈第2実施形態〉次に、本発明の第2実施形態を説明する。第2実施形態は、第1実施形態と共通点を有するので、相違点についてのみ説明する。第2実施形態は、本発明によるパケットの宛先情報管理装置を、バーチャル・プライベート・ネットワーク(VPN: Virtual Private Network)を構成するエッジルータに適用したものである。

【0113】VPNは、インターネット上で仮想的な私設ネットワークをつくる技術である。図10は、VPNの例を示す図である。図10に示すように、ISP(Internet Service Provider)のネットワーク51中に、各ユーザ(ここでは企業Aおよび企業B)が専用に利用できるVPN52、53が提供される。

【0114】企業Aの各サイト(端末装置)から送出されたパケットは、企業AのVPN52を通じてA社のみに伝わり、企業B等の他社に漏れない。また、企業AのVPN52には、企業B等の他社からのパケットも流れ込

まない。企業BのVPN53についてもVPN52と同様である。このような仮想的な専用線を、ISPのネットワーク51の中にいくつも共存させ、設備を共用することによって安価な接続を提供する。

【0115】通常、企業Aと企業Bの利用するアドレスは独立に決められるので、同じアドレスが使われる可能性がある。そのため、各エッジルータ54、55は、VPN52又はVPN53にパケットを転送する場合には、どの企業網のパケットかを示すVPN-idと呼ばれる識別子を、IPアドレスと共にパケットに付加する。

【0116】各エッジルータ54、55の構成は、図1に示したルータ1とほぼ同じ構成を有しているので説明を省略する。図11は、各エッジルータ54、55のルーティング機構3(図1参照)に設けられたルーティングテーブル20Bの例を示す図である。

【0117】図11に示すように、CAM6Dには、VPN-idとプレフィクスとの組み合わせで構成される拡張プレフィクスのエントリが格納されている。拡張プレフィクスは、VPN-idをなすビット列とプレフィクスをなすビット列とが直列に連結された連続するビット列である。

【0118】図11では、説明のため、“VPN-id: プレフィクス”を、“0:A”、“1:B”、“2:C”と表記しているが、実際には、可変長のビット列である。この例では、VPN-id=1は、企業AのVPN-idであり、VPN-id=2は、企業BのVPN-idである。さらに、VPN-id=0は、公衆インターネット用の識別子としての特殊なVPN識別子であり、VPNを使用せずにISPのネットワーク51を利用する者に対して割り当てられている。

【0119】以上の点を除き、CAM6D及びRAMツリー17Dの構成、経路作成部5による処理、アドレス検索部9によるルーティング処理は、第1実施形態と同じであるので説明を省略する。

【0120】第2実施形態によれば、CAM6Dのコンテンツとして、IPアドレスの他にVPN識別子(VPN-id)が追加された拡張プレフィクスが用いられ、VPNを意識することなくCAMエントリ及びRAMツリーが構成されている。これにより、CAM6Dへのアクセスの追加及びRAM7のデータ追加なしに、VPN毎にルーティング検索を分離することができる。

【0121】また、特別なVPN識別子(例えば、VPN-id=0)を公衆インターネット用の識別子とすることで、VPN及び非VPNのトラフィックを同じルーティングテーブル20Bで検索することができる。

【0122】〔付記〕本発明は、以下のように特定することができる。

(付記1) パケットの経路を選択するためのパケットの宛先情報を管理する装置であって、第1メモリと、前記

20

30

40

50

第 1 メモリよりも情報の検索速度が遅く且つ集積度が高い第 2 メモリと、前記宛先情報が前記第 1 メモリと前記第 2 メモリとに跨がる状態で格納された経路表を作成する作成部と、を備えたバケットの宛先情報管理装置。

(付記 2) 前記第 1 メモリの使用率の閾値を記憶した閾値記憶部をさらに備え、前記作成部は、前記第 1 メモリの使用率を取得し、取得した使用率が前記閾値記憶部に記憶された閾値未満である場合には、前記経路表に格納すべき宛先情報を前記第 1 メモリのみに格納する、付記 1 記載のバケットの宛先情報管理装置。

(付記 3) 前記経路表に格納された宛先情報に対応する出力ルートの情報を得る場合の最大許容時間を記憶した許容時間記憶部をさらに備え、前記作成部は、前記最大許容時間内で前記経路表から前記バケットの宛先情報が検索され、検索された宛先情報に対応する出力ルートの情報が取得されるように、前記宛先情報を前記第 1 メモリと前記第 2 メモリとに跨がる状態で格納する、付記 1 又は 2 記載のバケットの宛先情報管理装置。

(付記 4) 前記最大許容時間内で可能な前記第 2 メモリへの最大アクセス回数を記憶したアクセス回数記憶部をさらに備え、前記作成部は、前記経路表に格納された宛先情報の検索に必要な前記第 2 メモリへのアクセス回数が前記最大アクセス回数以上となる場合には、前記アクセス回数が前記最大アクセス回数未満となるように、前記第 2 メモリに格納された、又は格納される予定の宛先情報を前記第 1 メモリに格納する付記 3 記載のバケットの宛先情報管理装置。

(付記 5) 前記宛先情報は、所定長のビット列で構成され、前記作成部は、前記ビット列の最上位ビットから所定ビットまでを前記第 1 メモリに格納し、前記所定ビットから最下位ビットまでを前記第 2 メモリに格納する、付記 1 記載のバケットの宛先情報管理装置。

(付記 6) 前記宛先情報は、バケットの宛先 IP ネットワークアドレスである、付記 1 記載のバケットの宛先情報管理装置。

(付記 7) 前記宛先情報は、バーチャル・プライベート・ネットワークの識別子と、バケットの宛先 IP ネットワークアドレスとの組み合わせである、付記 1 記載のバケットの宛先情報管理装置。

(付記 8) 前記作成部は、前記識別子を前記第 1 メモリのみに格納する、付記 7 記載のバケットの宛先情報管理装置。

(付記 9) バケットの宛先情報が第 1 メモリとこの第 1 のメモリよりも情報の検索速度が遅く且つ集積度が高い第 2 メモリとに跨る状態で格納された経路表と、入力されたバケットに対応する宛先情報を前記経路表から検索し、検索した宛先情報に対応する出力ルートの情報を取得する検索部と、を備えたバケットの経路選択装置。

(付記 10) バケットの経路を選択するためのバケットの宛先情報を管理装置によって管理する方法であって、

前記管理装置は、前記宛先情報が第 1 メモリとこの第 1 メモリよりも情報の検索速度が遅く且つ集積度が高い第 2 メモリとに跨る状態で格納された経路表を作成する、ことを含むバケットの宛先情報管理方法。

(付記 11) 前記管理装置は、前記第 1 メモリの使用率の閾値を記憶し、記憶された前記第 1 メモリの使用率を取得し、取得した使用率が前記閾値記憶部に記憶された閾値未満である場合には、前記経路表に格納すべき宛先情報を前記第 1 メモリのみに格納する、付記 10 記載のバケットの宛先情報管理方法。

(付記 12) 前記管理装置は、前記経路表に格納された宛先情報に対応する出力ルートの情報を得る場合の最大許容時間を記憶し、前記最大許容時間内で前記経路表から前記バケットの宛先情報を検索し、検索した宛先情報に対応する出力ルートの情報が取得されるように、前記宛先情報を前記第 1 メモリと前記第 2 メモリとに跨る状態で格納する、付記 10 又は 11 記載のバケットの宛先情報管理方法。

(付記 13) 前記管理装置は、前記最大許容時間内で可能な前記第 2 メモリへの最大アクセス回数を記憶し、前記経路表に格納された宛先情報の検索に必要な前記第 2 メモリへのアクセス回数が前記最大アクセス回数以上となる場合には、前記アクセス回数が前記最大アクセス回数未満となるように、前記第 2 メモリに格納された、又は格納される予定の宛先情報を前記第 1 メモリに格納する付記 12 記載のバケットの宛先情報管理方法。

(付記 14) 前記宛先情報は、所定長のビット列で構成され、前記管理装置は、前記ビット列の最上位ビットから所定ビットまでを前記第 1 メモリに格納し、前記所定ビットから最下位ビットまでを前記第 2 メモリに格納する、付記 10 記載のバケットの宛先情報管理方法。

(付記 15) 前記宛先情報は、バケットの宛先 IP ネットワークアドレスである、付記 10 記載のバケットの宛先情報管理方法。

(付記 16) 前記宛先情報は、バーチャル・プライベート・ネットワークの識別子と、バケットの宛先 IP ネットワークアドレスとの組み合わせである、付記 10 記載のバケットの宛先情報管理方法。

(付記 17) 前記管理装置は、前記識別子を前記第 1 メモリのみに格納する、付記 16 記載のバケットの宛先情報管理方法。

(付記 18) バケットの経路選択装置によってバケットの経路を選択する方法であって、前記経路選択装置は、バケットの宛先情報が第 1 メモリとこの第 1 のメモリよりも情報の検索速度が遅く且つ集積度が高い第 2 メモリとに跨る状態で格納された経路表を作成し、入力されたバケットに対応する宛先情報を前記経路表から検索し、検索した宛先情報に対応する出力ルートの情報を取得する、ことを含むバケットの経路選択方法。

[0 1 2 3]

【発明の効果】本発明によれば、CAM方式で作成されたルーティングテーブルよりも多くのプレフィックスを保持することができ、且つRAMツリー方式で作成されたルーティングテーブルよりも高速にプレフィックスを検索することができる。

【図面の簡単な説明】

【図 1】 ルータの構成図

【図 2】 CAMの構成図

【図 3】 ノード(エントリ)の追加処理の説明図

【図 4】 ルーティングテーブルの説明図

【図 5】 プレフィックス追加の処理を示すフローチャート

【図 6】 プレフィックス追加の処理を示すフローチャート

【図 7】 RAMツリーの段数を減らす処理の説明図

【図 8】 CAMエントリ削除の例を示す図

【図 9】 RAMエントリ削除の例を示す図

【図 10】 VPNの説明図

【図 11】 ルーティングテーブルの説明図

【図 12】 従来技術の説明図

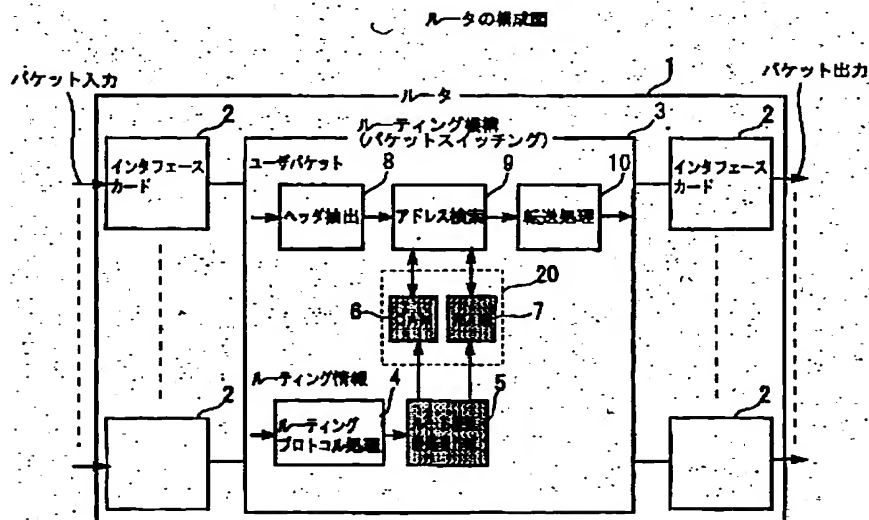
【図 13】 従来技術の説明図

【図 14】 従来技術の説明図

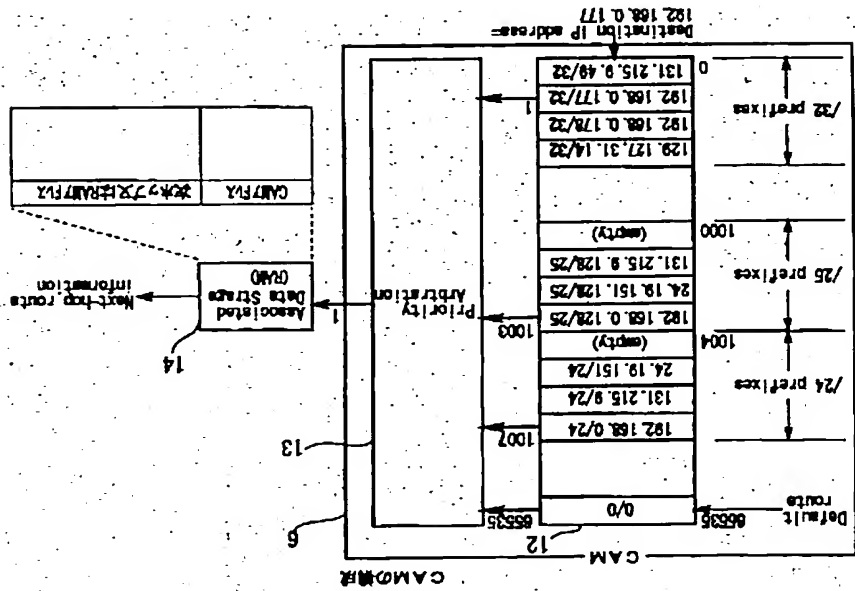
【符号の説明】

- 1 ルータ
- 2 インターフェイスカード
- 3 ルーティング機構
- 4 ルーティングプロトコル処理部
- 5 ルート計算・経路表作成部
- 6 CAM
- 7 RAM
- 10 転送処理部
- 8 ヘッダ抽出部
- 9 アドレス検索部
- 12 データアレイ/マスクアレイ
- 13 プライオリティ・エンコーダ
- 14 補助RAM
- 17 RAMツリー
- 20 ルーティングテーブル

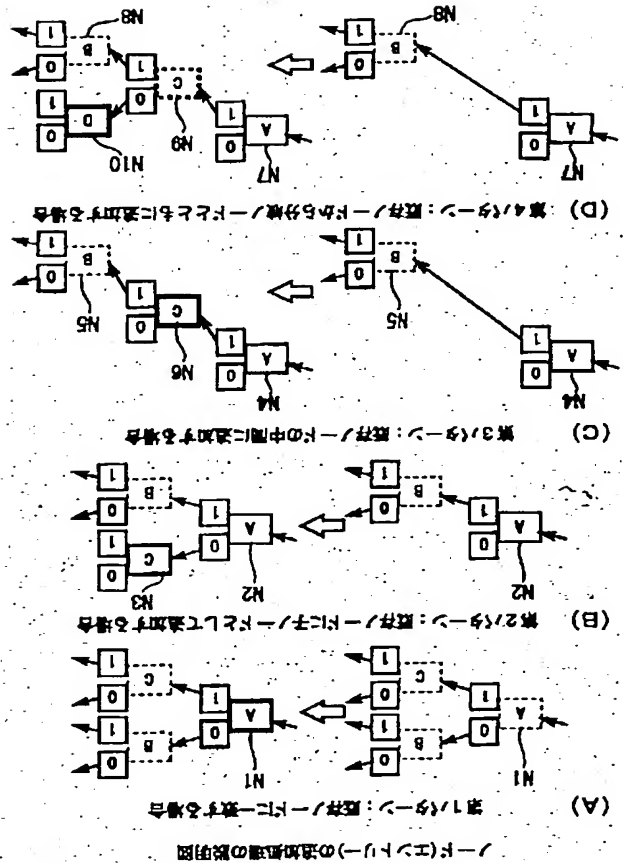
【図 1】



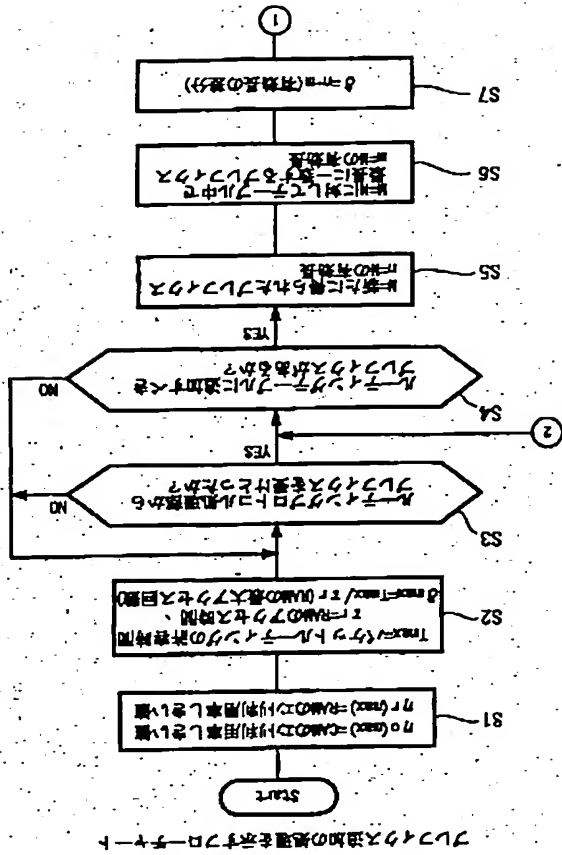
【図 2】



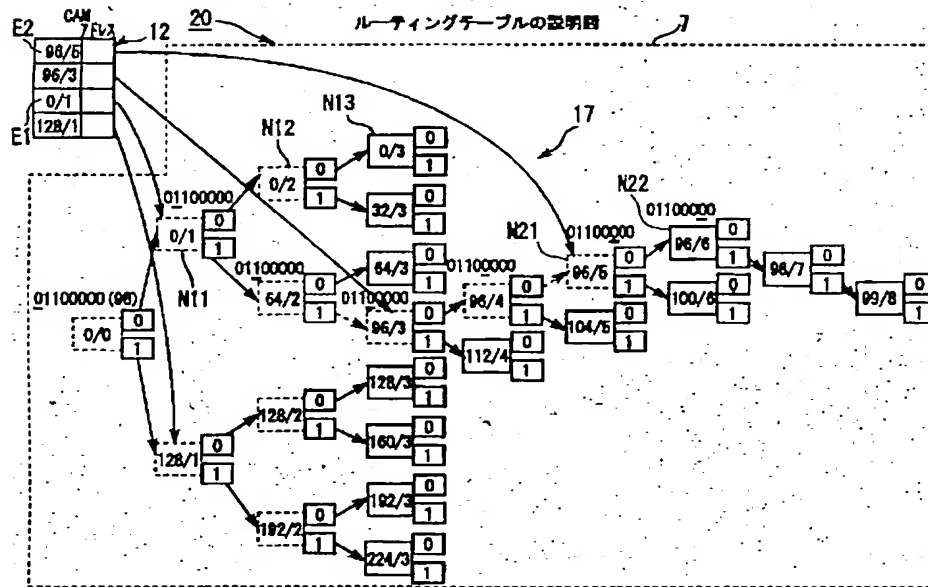
【図 3】



【図 5】



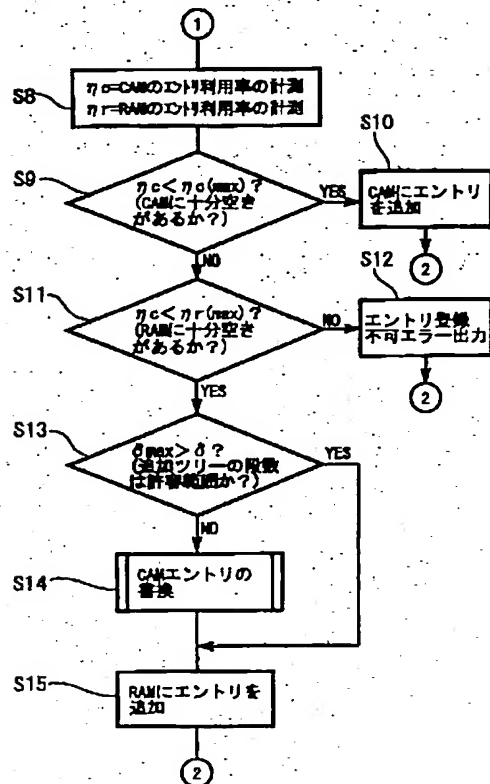
【図4】



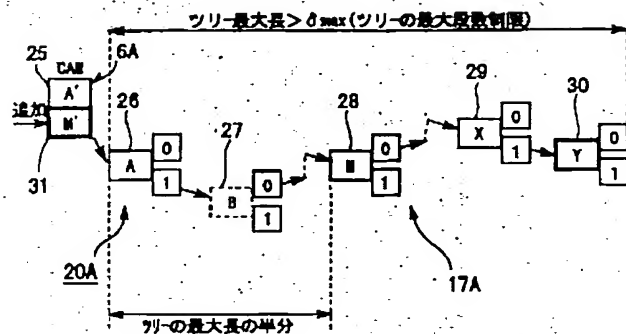
【図6】

【図7】

プレフィクス追加の処理を示すフローチャート

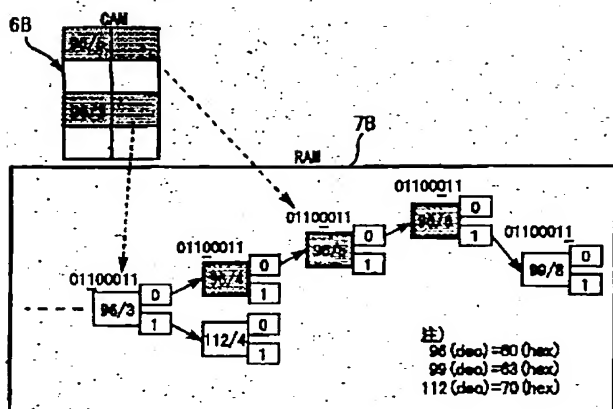


RAMツリーの段数を減らす処理の説明図



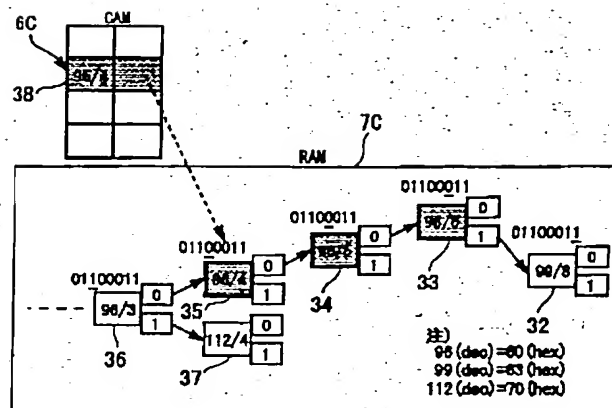
【図 8】

CAM エントリ削除の例を示す図



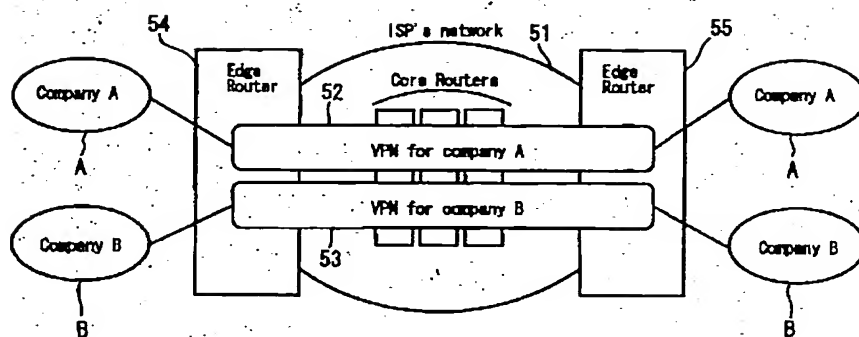
【図 9】

RAM エントリ削除の例を示す図

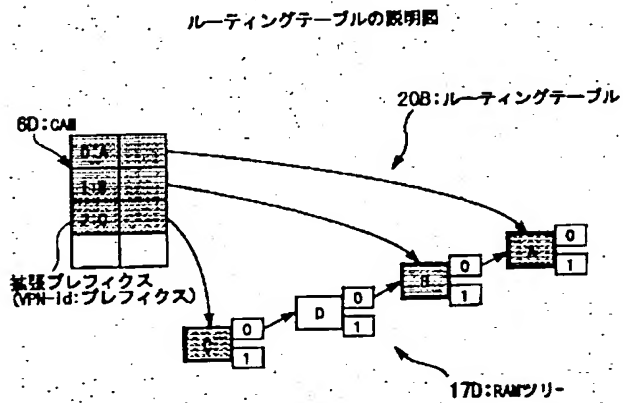


【図 10】

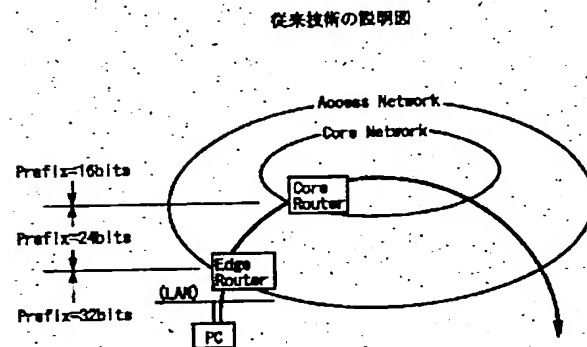
VPN の説明図



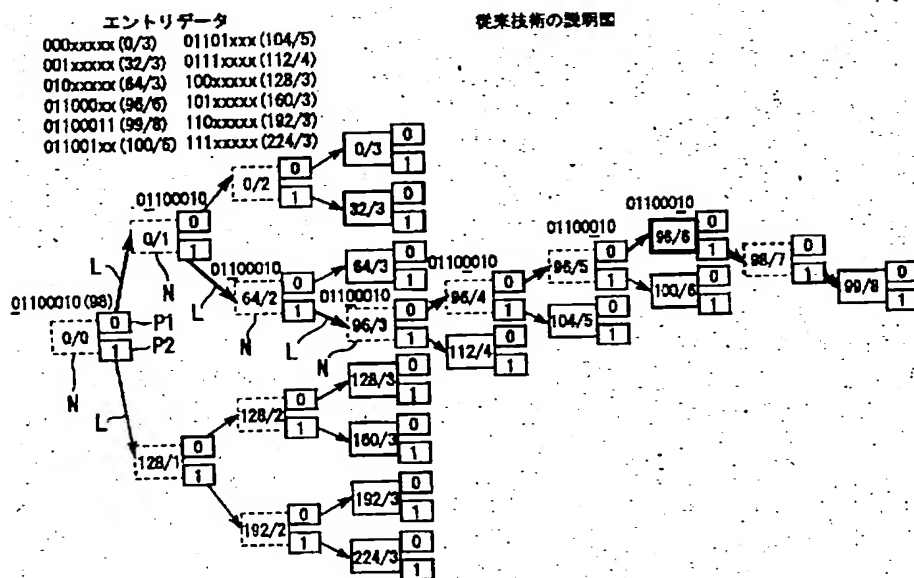
【図 11】



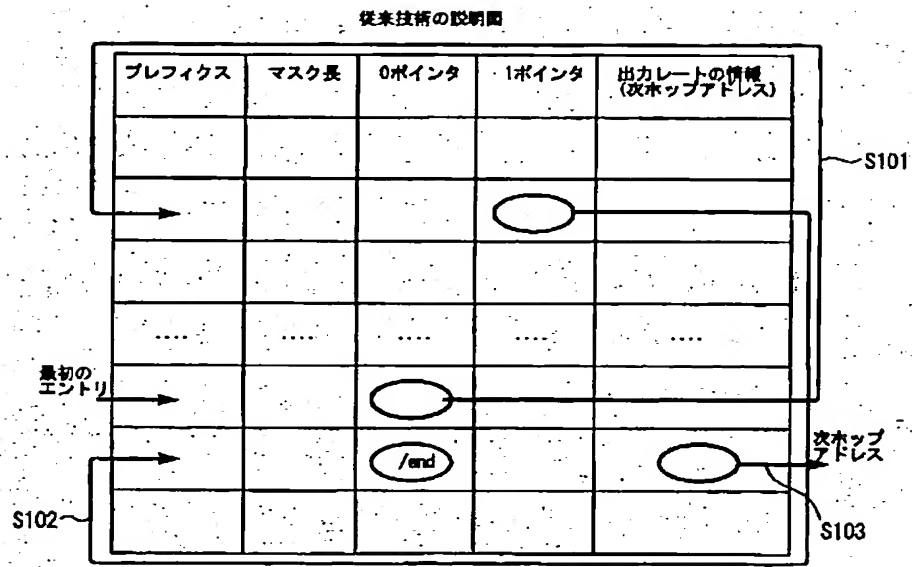
【図 12】



【図 13】



【図 14】



フロントページの続き

Fターム(参考) 5K030 GA01 GA06 HA08 HB29 HD03
 HD09 KA01 KA05 KA13 LB05
 LE03 MA13